ESC-TR-92-121

Annual Technical Summary

Parametric Study of Diffusion-Enhancement Networks for Spatiotemporal Grouping in Real-Time Artificial Vision

R.K. Cunningham
A.M. Waxman

24 July 1992

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS

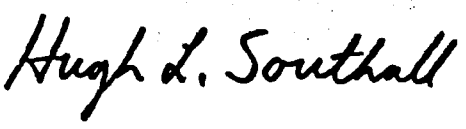
Prepared for the Department of the Air Force under Contract F19628-90-C-0002.

This report is based on studies performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology. The work was sponsored by the Air Force Office of Scientific Research under Contract F19628-90-C-0002.

The ESC Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

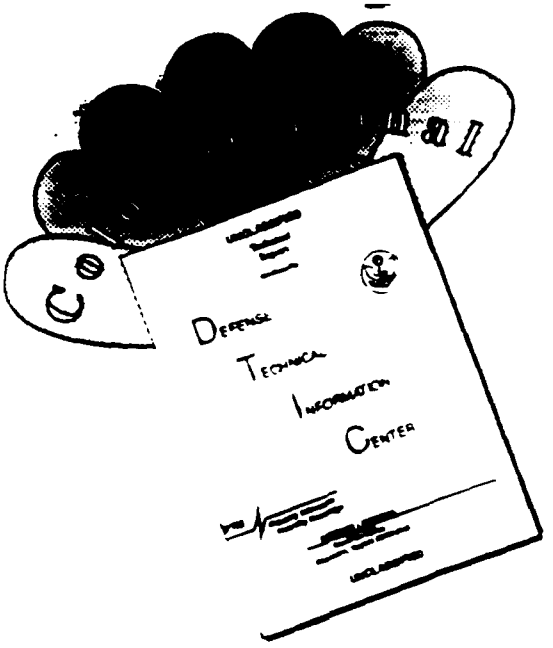
This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER



Hugh L. Southall, Lt. Col., USAF
Chief, ESC Lincoln Laboratory Project Office

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF COLOR PAGES WHICH DO NOT REPRODUCE LEGIBLY ON BLACK AND WHITE MICROFICHE.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

PARAMETRIC STUDY OF DIFFUSION-ENHANCEMENT NETWORKS FOR SPATIOTEMPORAL GROUPING IN REAL-TIME ARTIFICIAL VISION

ANNUAL TECHNICAL SUMMARY REPORT
TO THE
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

APRIL 1991 — JUNE 1992

R.K. CUNNINGHAM
A.M. WAXMAN
Group 21

24 JULY 1992

Approved for public release: distribution is unlimited.

LEXINGTON MASSACHUSETTS

ABSTRACT

This is the second Annual Technical Summary of the MIT Lincoln Laboratory parametric study of diffusion-enhancement networks for spatiotemporal grouping in real-time artificial vision. Spatiotemporal grouping phenomena are examined in the context of static and time-varying imagery. Dynamics that exhibit static feature grouping on multiple scales as a function of time, and long-range apparent motion between time-varying inputs, are developed for a biologically plausible diffusion-enhancement bilayer. The architecture consists of a diffusion layer and a contrast-enhancement layer coupled by feedforward and feedback connections; input is provided by a separate feature-extracting layer. The model is cast as an analog circuit that is realizable in VLSI, the parameters of which are selected to satisfy a psychophysical data base on apparent motion.

PREFACE

This report contains the selected results of many simulations performed during the past year using the network architecture and simulation tools developed during the first contract year [4]. Although all these simulations are new this year, the underlying biological and psychophysical insights and motivations remain the same; thus, the sections describing these findings are similar to the first year's report. Small additions and corrections have been added where new results have come to light, a more detailed description of a possible cortical location for this processing is indicated, and a number of the figures have been reworked.

In addition, though few significant changes have been made to the network, many of the perceptual effects described below have now been demonstrated. This success is due to the change in approach discussed at the end of last year's report [4], in which the "parameter pools" that produce the desired effects are located by human-guided search. As a result, the section describing the model is also quite similar to last year's, though some changes have been made to the feedback characterization.

Those familiar with the first-year report might wish to skim these sections and start reading in detail at the section entitled "Diffusion-Enhancement Bilayer Model."

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

1. INTRODUCTION

This study suggests that diffusion-enhancement interactions play a fundamental role in human preattentive perception. Understanding parallel networks that simulate such interactions is important for understanding neurobiological findings and will suggest new experiments for researchers in that field. The study of spatiotemporal networks is of special interest because it aids the understanding of the static and dynamic grouping of stimuli that occurs in and across the visual, auditory, and somatosensory systems [14,24,25].

Such grouping phenomena are also important in artificial vision systems [19]. During the past year, a biologically plausible network model has been refined [2,3] from several reference works [4,19,21,22,23] and used to explain some static and dynamic grouping experiments from the psychophysical literature. This diffusion-enhancement bilayer (DEB) represents the interactions between a layer of astrocyte glial cells and a neuronal layer. Long-range communication is achieved via the diffusion of K^+ throughout the electrically coupled glial layer, and percept localization occurs in the contrast-enhancing (CE) neuronal layer. These glial and neuronal layers are coupled by K^+ currents leaking from glial endfeet in close proximity to the neuronal layer.

In order to explain the DEB model, the reader must understand some of the biological and psychophysical results that led to the development of this model; thus, a literature review is presented before the model itself is described. Numerical simulations performed with the model are also discussed.

During the last year, two feedback parameterizations were considered; both are reported here. The first is a simple algebraic decay that has been used to demonstrate several psychophysical phenomena. This parameterization is then shown to be sensitive to system noise. As a result, a second parameterization is considered in which the feedback activates more slowly. This is significantly more resistant to noise, but has not yet been shown to demonstrate all of the modeled phenomena. Completed simulations are then presented and plans for future simulations outlined.

2. BACKGROUND: PSYCHOPHYSICAL PHENOMENA AND RELATED NEUROBIOLOGY

Psychophysical experiments that are being modeled and neurobiology that could support such phenomena are described in this section. Both psychophysics and neurobiology provide data used to set the DEB model parameters.

2.1 Psychophysical Results

2.1.1 Spatiotemporal Grouping Phenomena

The DEB model replicates the effects of both static and dynamic grouping. A striking example of static feature grouping in the visual domain is demonstrated by Marroquin's diagram, shown in Figure 1 [17]. Notice how the dots in the diagram appear to group with their neighbors on greater and greater scales over time as one stares at the center of the hexagon.

The simplest dynamic apparent motion effect is known as gamma motion and occurs when a single light is turned on for a brief time, then turned off. Although the light is of a fixed spatial extent, the perception is of a light that first expands and then contracts [14], as shown in Figure 2 for one-dimensional space.

New dynamic grouping phenomena emerge when two distinct stimuli interact over time to form the percept of long-range apparent motion. In the human visual system, apparent motion can be demonstrated with two lights of fixed spatial extent that are illuminated at distinct times across a fixed spatial separation (Figure 3). With different spatial separations, illumination times, and interstimulus intervals, the light can appear as two separate lights flashing, as one spot that moves smoothly between two real lights, or as one spot that moves smoothly from the first location, jumps, and continues moving smoothly to the second location [14,24]. The most common form of object motion, phi, gives no impression of a particular shape undergoing motion, whereas in beta motion a well-defined shape is seen in motion. Similar effects can be achieved by substituting tones for points of light [25].

Psychologists have examined in detail the conditions that will produce these three distinct types of motion, and discovered that for fixed-flash durations there is a clear range of onset-to-onset interval (SOA)[1] versus spatial separation that will produce smooth apparent motion (Figure 4). If the SOA is shortened, lights begin jumping rather than smoothly moving between each other, while

[1]The onset-to-onset interval is often referred to as the "stimulus onset asynchrony" (SOA) and is defined as the time between the onset of two successive applications of a stimulus. Related is the interstimulus interval (ISI) that is defined as the time between two successive stimuli. Thus, ISI + stimulus duration = SOA.

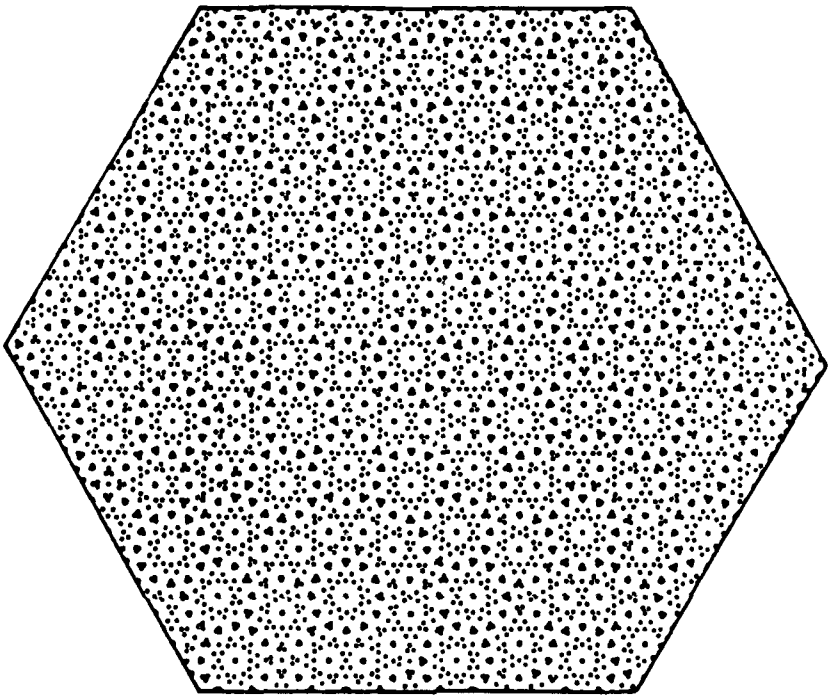


Figure 1. Static feature grouping on multiple scales: evidence for the existence of a dynamic grouping process in the visual system [17].

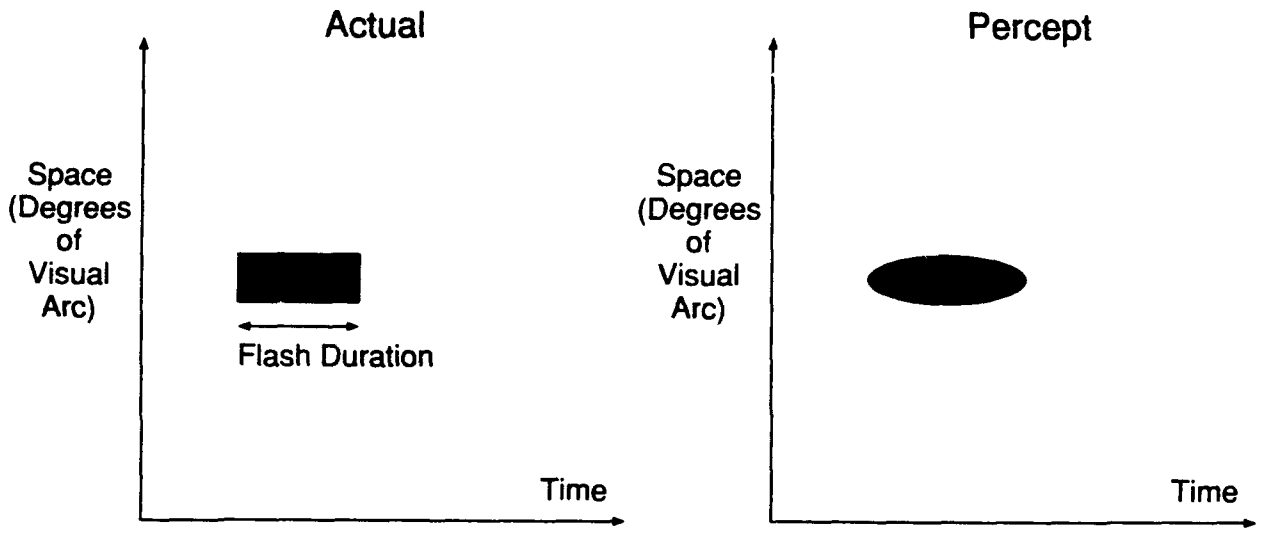


Figure 2. Gamma motion. Left: a light of fixed spatial extent is illuminated then extinguished. Right: the percept is of a light expanding and then contracting.

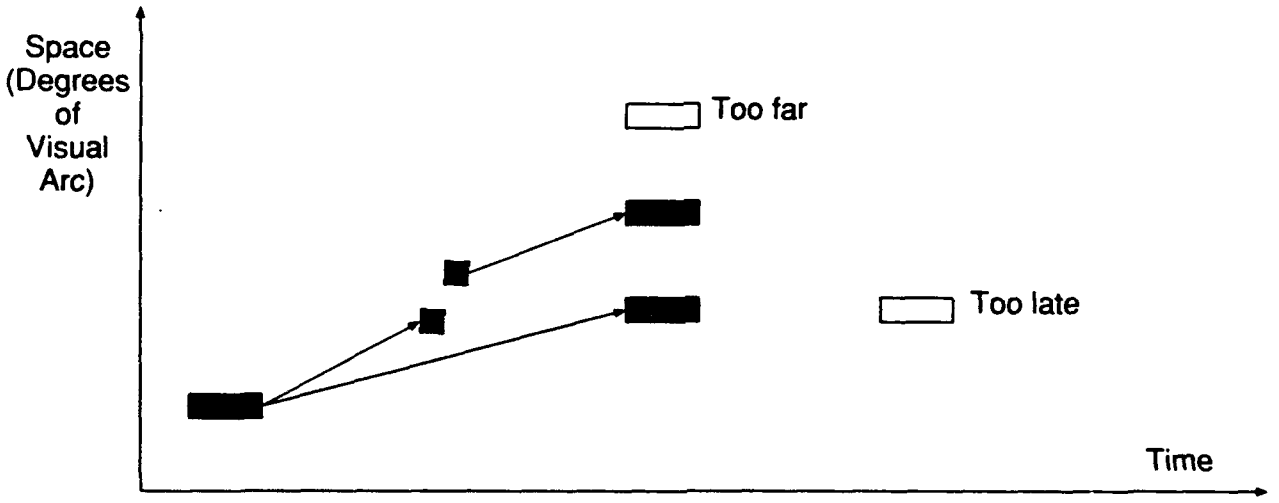


Figure 3. Long-range apparent motion. The filled rectangles represent sources in space-time separations that produce the illusion of long-range apparent motion. The empty rectangles represent sources that are ignited too late for the spatial distance or are too far away for the given ISI. The bottom center rectangle occurs soon enough to give the appearance of smooth motion from the left-most rectangle to the first shaded square, a short jump to the second shaded square, followed by smooth motion to the destination, while the rectangle below it exhibits pure smooth motion.

a still shorter SOA causes the lights to appear to flash simultaneously. If the SOA is lengthened beyond an acceptable limit, the lights flash independently of each other. Similar conditions can be created by varying the spatial separation of the two stimuli.

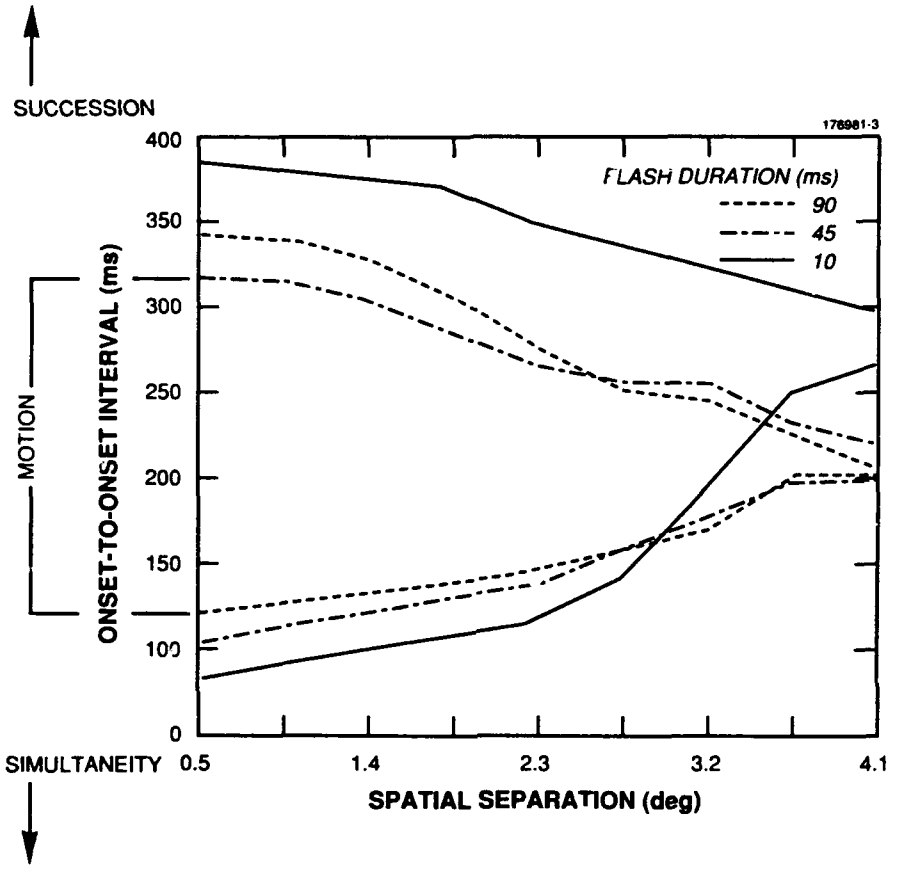


Figure 4. Parameters affecting long-range apparent motion, SOA versus spatial separation [14].

Once this effect was discovered, the next step was to calculate rate of motion. Unfortunately, it is unclear how best to calculate velocity of the illusory motion. A physical theory would plot $V = \frac{\text{Distance Traveled}}{\text{Time}}$; in these experiments the distance traveled is clear, but the interpretation of time is not. If ISI is used (the time between when the object was last seen in its initial location and when it showed up in its final location), cases exist for illusory motion with infinite velocity. SOA is the next reasonable choice, but stimulus duration has been shown to affect apparent motion. Kolers [14] argues that this is as reasonable a plot as any; this report agrees and displays velocity as distance traveled versus SOA in Figure 5 for comparison purposes. It is unclear, however, whether the relationship between velocity and spatial separation is linear.

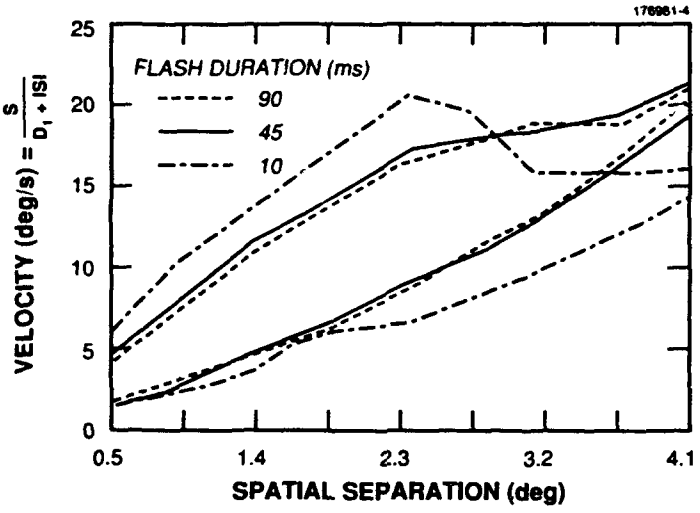


Figure 5. Calculated velocity versus spatial separation [14].

To further understand apparent motion, other psychophysical experiments have been performed; in the visual domain all involve additional stimuli. *Split-motion* effects are demonstrated with three lights of fixed spatial extent; the center light is illuminated and extinguished, then the two outer lights. If the two outer lights are equidistant to the center light, the latter appears to split and move to both outer lights. If the two outer lights are at staggered distances, the center light appears to move only to the closest light (Figure 6). Also, if the equidistant version of the display is placed in the periphery of the visual field, the center light appears to move toward the light farthest from the fovea.

The opposite of the split effect is the *merge effect*, in which first the two outer lights, then the center light, are illuminated and extinguished. If the two outer lights are equidistant to the center, they both appear to move to the center light and merge with it. If the two outer lights are not equidistant, only the one that is closer appears to move to the center (Figure 7). Once again, if the equidistant display is placed in the periphery of the visual field, the outermost light appears to move toward the center.

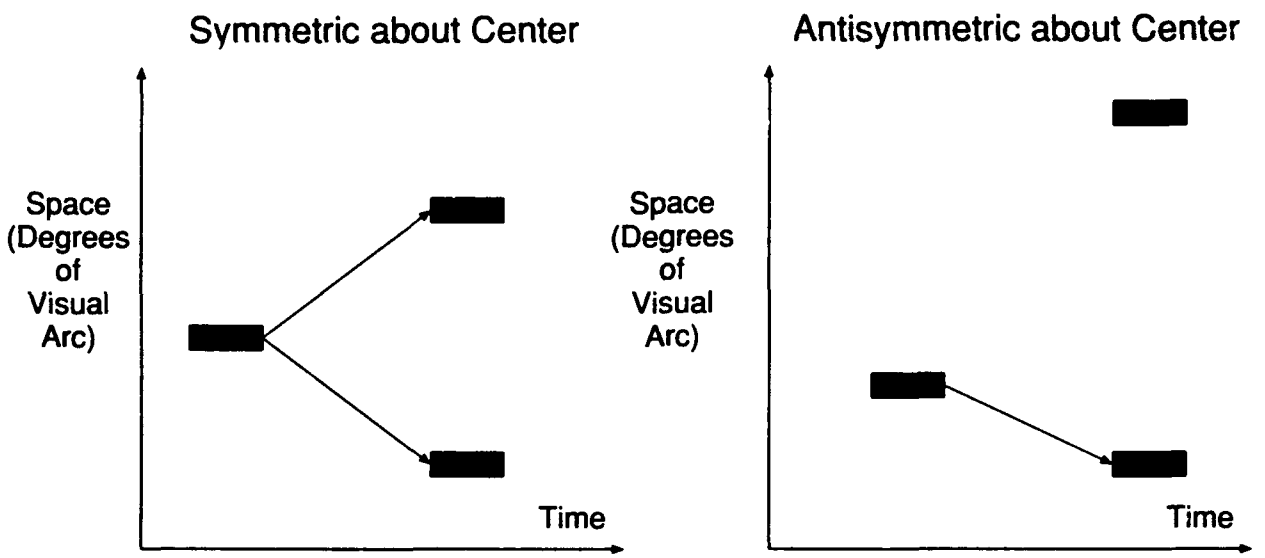


Figure 6. Split effects. Illuminating one point source followed by illuminating two equidistant point sources cause the first to appear to split and move to both of the later sources. If these are not equidistant to the first, movement is only to the closer of the pair.

Another important multielement stimulus causes the Ternus effect. This effect illustrates that more than one distinct motion percept can be achieved from the same display by altering subtle aspects. In the Ternus stimulus, two frames with three aligned lights each are illuminated and extinguished in succession. The frames are aligned so that two of the three lights occupy the same space, and the third appears alternately on the left and right of the central two objects. When the ISI is short, the third light appears to move around the central two objects (element motion), but when the ISI is long the three appear to shift as a coherent group (Figure 8).

Although the cause of these motion effects has been debated for nearly a century, it is known that they do not occur at the retinal level, as evident from a variation on the basic long-range apparent motion experiment. In this variant, known as dichotic presentations, the first light is presented to the subject's left eye, and the second light is presented to the subject's right eye. Apparent motion is experienced, indicating that spatiotemporal interactions occur at the cortical level. In addition, and based on the split and merge experiments performed in the periphery, these effects occur after the visual system has compressed (or down-sampled) the periphery in favor of the fovea, i.e., at the visual cortex. Such sampling and compression occur in at least two places in the visual system [11]. First, the retina itself contains a nonuniform population of rods and cones; the densest population occurs in the fovea. Second, the receptive fields of the retinal ganglion cells are markedly smaller in the fovea than in the periphery; hence, the well-known cortical magnification

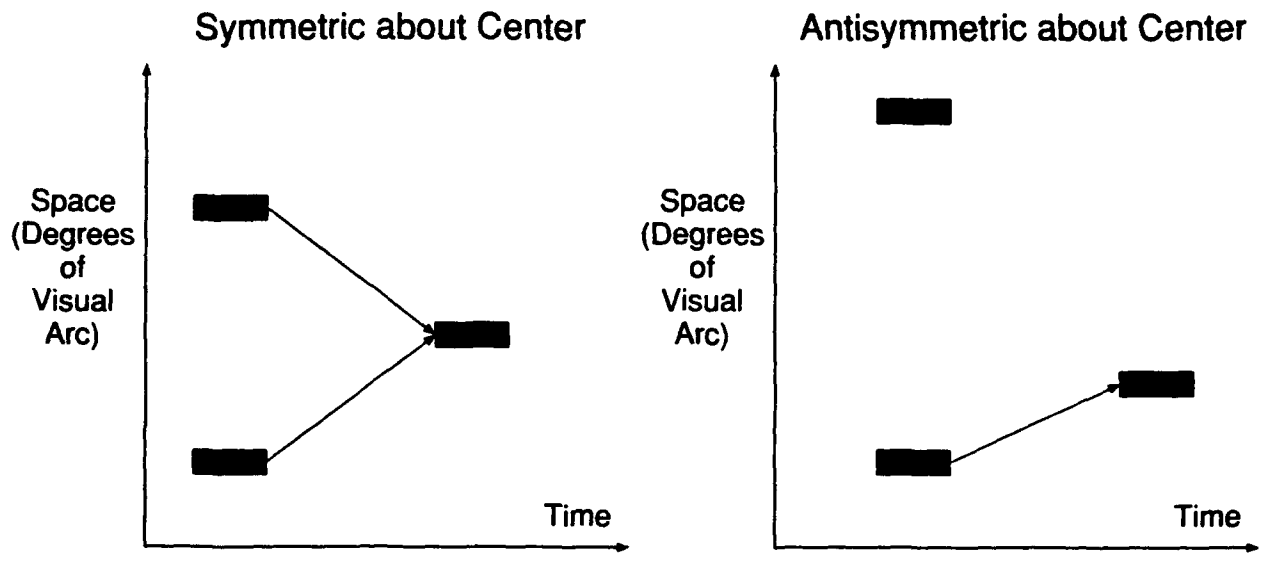


Figure 7. Merge effects. Illuminating two point sources equidistant from a third, later-illuminated point source gives the appearance of the two merging and becoming one. If the original pair is not equidistant from the third point source, only the closer of the two will appear to move.

of the fovea and compression of the periphery. Based on neurobiology and psychophysics, it seems evident that the substrate of apparent motion lies in the visual cortex.

This grouping process is not restricted to sight. In the tactile area, if two vibrators agitate the skin with a small ISI, the subject experiences a single vibration between them. As in the visual domain, the effect occurs not in the skin but in the cortex; if the skin between the two vibrators is locally anaesthetized, the effect is still experienced [7]. Similar experiments have been devised for the auditory sense, and similar results were reported. Most bizarre are the intermodal experiments, in which apparent motion is perceived between a sound and a light source [25].

2.1.2 Useful Psychophysical Parameters

Psychophysical literature suggests that the structure of the DEB model and long-range apparent motion examples continue to support this diffusion concept and begins to suggest rates at which it should occur. Figure 4 provides spatial and temporal parameters to which the model should conform. The lower curves suggest that communication time varies with distance (i.e., Korte's Law), while the upper curves suggest that object memory (and its reinforcement mechanisms) will fade after a time and that this fade occurs faster over greater distances. Both are considered indicative of a leaky diffusion network in which stimulus activity first rises and spatially expands, then falls

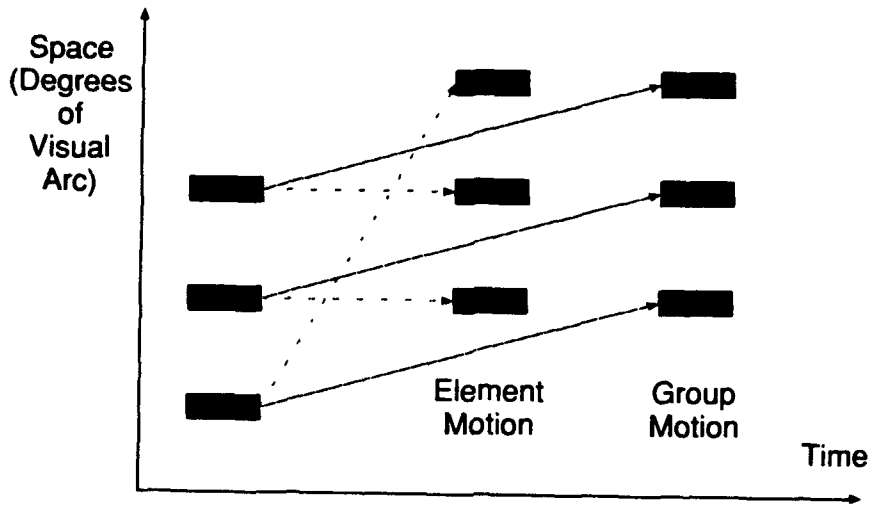


Figure 8. Ternus effect. Illuminating shifting groups of three point sources produces the illusion of one in element motion or all three in group motion. The percept changes with changing ISI.

and contracts; the DEB network was designed to produce these effects. For a fixed (small) spatial separation, there should be a range of SOAs in which motion should be perceived; outside that range no motion should be realized.

It is difficult to directly relate the model time and space scales to biological time and space, so ratios are examined that allow units to be ignored. One ratio considered is the longest-to-shortest SOA for which smooth motion occurs at a fixed spatial separation. Figure 4 suggests that this ratio is restricted to ≤ 4, with the ratio decreasing as the spatial separation increases. It is uncertain how to compare psychophysical with model-predicted velocities because there is still no way to interpret the velocities of apparent motion.

2.2 Biological Considerations

Consider that the above psychophysical effects have two salient processes: long-range communication that facilitates interaction of features generated by the inputs (point light sources in the visual domain) and a focusing process that enables objects in apparent motion to have a definitive location. This section discusses the neurobiological elements that are most likely to provide the foundations for both processes. Because it has already been suggested that static and dynamic grouping occurs in the cortex, the cortex is seen as the likely host.

The cortex is divided almost evenly between nonneuronal and neuronal cells. The largest class of nonneural brain cells, the neuroglia, is believed to be the site of the long-range communication process, while neuronal networks provide focusing of activity to create the percept.

There are strong indications from psychological, neurophysiological, and perceptual literature that two rather separate pathways exist; one that responds to movement, and a second that responds to static attributes such as form and color [16,26]. The dynamic pathway is commonly referred to as the magnocellular pathway. In primates, the anatomical and physiological differentiation from the static or parvocellular pathway is clear as early as the retinal ganglion cells. Large, type-A retinal ganglion cells provide input to the large-celled magnocellular layers of the lateral geniculate body (LGN), while smaller, type-B cells provide input to the parvocellular subdivision of the lateral geniculate body. From the LGN, the magnocellular pathway progresses up through visual cortex areas V1 layers 4Cα into 4B; then into areas V3, V2K, and V5 simultaneously. The parvocellular pathway enters 4A both directly and via 4Cβ, then enters layers 2 and 3 before connecting with areas V2I and V2N, which in turn connect to V4. After this point connections are more broadly made and less well understood. Cortical areas V2 and V3 could be the site of the processing illustrated in this report, and the two pathways permit a similar network to be used in both.

2.2.1 Astrocyte Glial Cells

Once thought of as only providing passive physical support, neuroglial cells now appear to play an "active role in maintaining normal brain physiology" [13]. Concentration is on the astrocyte glial cells because they are known to provide long-range communication between coupled astrocytes (Figure 9). Although to date such coupling has not been directly demonstrated in vivo (Kettenman and Ransom [12] suggest that this is due to technical difficulties), there is some evidence that it occurs,[2] and there is direct evidence for coupling in cultured astrocyte cells. Such communication is not rare. Indeed, Kettenman and others have observed that "mammalian astrocytes in cell culture are widely coupled to one another electrically" and that "qualitative studies have shown that cultured astrocytes form a highly coupled electrical syncytium" [12] that is believed to provide the long-range communication necessary to support the above psychophysical phenomena.

2.2.2 Neuronal Networks

Many cells in the visual cortex are known to derive their input from networks of neurons. In Hubel [11], simple cells (which respond to oriented lines) are postulated to be made up of a hierarchy of lower-order, radially symmetric, center-surround cells. Similarly, complex cells (which respond to oriented lines in a wide receptive field) and end-stopped cells are made up of a network of simple cells. Also, directionally tuned, motion-sensitive cells are postulated to consist of inhibitory

[2]Low-molecular-weight dye passes between adjacent cells [1], and glial networks are postulated to act as potassium spatial buffers [12,13,15,18].

and excitatory connections. Similar inhibitory and excitatory interconnections are used in the CE neuronal network where cells are assumed to be self-exciting and latterly inhibiting within the layer.

2.2.3 Interactions

Astrocyte glial cells are known to interact electrically among themselves and with neurons; the nature of that interaction is now considered.

Kettenman and Ransom [12] discovered that in cultured astrocyte syncytia, the resistance of the electrical gap junctions between astrocytes is not voltage dependent over much of the membrane potential fluctuation; thus, a charge flow model examining intra-astrocyte communication should be independent of the membrane potential of the astrocytes. Futhermore, it is known that "glial cells...have a high potassium concentration and have negligible ionic permeability for ions other than potassium" [15]; therefore, current flow is modeled through glial cells as transfer of potassium ions to or from the cell in a manner obeying Ohm's law. When this process is expanded to encompass current flow in a network of glial cells, the motion of these ions can be approximated with a diffusion equation.

A further proposition is that a glial syncytia provides long-range communication between neurons in a layer via transmission of potassium and other ions. It has been shown that "variations in $[K^+]_{\text{extracellular}}$ have profound effects on neuronal excitability, modulating such processes as synaptic transmission and the initiation and propagation of action potentials" [18]. Such $[K^+]$ variations can be realized near the leaky endfeet of astrocytes that are in close proximity to neuronal synapses (see Figure 9). This report is not the first to propose such an interaction. In 1965, Hertz [10] proposed "a mechanism...in which the potassium ions, which have been lost from one nerve cell during its activity, are transported through neuroglia cells to the outer surface of another nerve cell, which is then depolarized and stimulated; that is, a neuronal-neuroglial-neuronal impulse transmission." Hertz continues: "Potassium ions which have been released from an active area are transported through neuroglia cells to the outside of other neurones [sic]; these are in turn stimulated and potassium ions are released, to be transported actively through other neuroglia cells. In this way the spreading depression is propagated across the entire cortex more rapidly than can be explained by a diffusion." An alternative rapid propagation mechanism could be due to small interglial electromagnetic fields, where ions taken up at one location induce other ions to be released elsewhere. The DEB model explicitly uses such interactions to spread and reinforce the charge distribution in a diffusion layer.

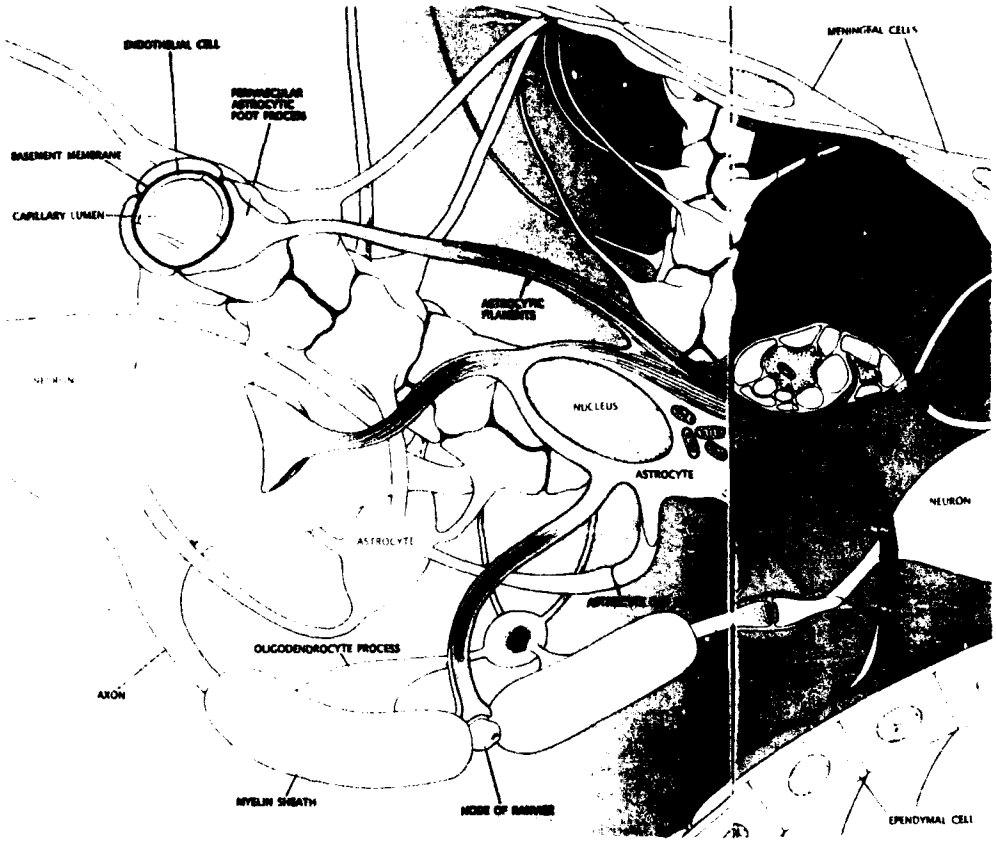


Figure 9. Neuroglial-neuronal interactions. Astrocytes are present near specific synapses to take up and metabolize specific neurotransmitters [13], as well as mediate K^+ ion concentrations [15,18]

2.2.4 Useful Biological Parameters

Odette and Newman [18] note that glial cell endfeet "can contain up to 95% of the total cell conductance." This is important to determine how "leaky" the diffusion process should be.

Kettenman and Ransom [12] have examined astrocyte coupling in cultured layers by electrically stimulating (via KCl injections) one glial cell, then measuring its voltage and that of the neighboring cell. The ratio of these voltages is fit to an exponential, which approximates the steady-state decay in a 1-D and 2-D syncytium: $\frac{V_2}{V_1} = \exp\left(\frac{-d}{L}\right)$, where d is the distance from the injection and L is the length constant. Kettenman and Ransom [12] measured astrocyte L in vitro to be 80 to 100 μm. L can be used to relate the DEB model to physical size of the biological networks and is related to the ratio of the conductances given by G_{gg} and G_g in Section 3.1. Further, L is not related to the 1-D model explained herein, as it is believed that the 1-D length constant would have to be significantly greater than the 2-D decay length. Indeed, experiments with restricted 2-D syncytia have L values that are greater than their full 2-D syncytia counterparts [12]. L is expected to be more valuable in the context of the 2-D experiments.

3. DIFFUSION-ENHANCEMENT BILAYER MODEL

3.1 DEB Network Architecture

The DEB model consists of two processes that mirror the two salient psychophysical processes mentioned in Section 2.2, i.e., a diffusion layer that facilitates long-range interactions via local connections and a focusing layer that reinforces the diffusion layer and provides the sensation of a localized object traversing a spatial separation. In this model, visual input is presented and preprocessed to point features before passing to the spatiotemporal grouping network. In the case of the primate vision system, both center-surround processing and logarithmic spatial mapping occur before grouping begins in the cortex [11]. Following feature extraction, activity is input to the diffusion layer that interacts with a localizing CE layer, which periodically samples the state of the diffusion layer; its output is fed back to the diffusion layer to reinforce new input and facilitate sustained interactions. This report proposes that a motion detection system, such as that of van Santen [20] or Waxman [22] or [6], detects the smooth motion of the activity maximum as well as the motion of the activity edges at the output of the CE layer and causes the sensation of motion in the psychophysical experiments. Also, activity prompted by a single input at first grows, then eventually dies down, so that after a period of time grouping is no longer possible (Figure 4). This effect is a result of the limited time span of featural input from a single feature, the leaky diffusion layer, and the imposition of decay on the feedback from the CE layer.

With this high-level description of the network in mind, a 1-D circuit form of the DEB model is illustrated in Figure 10. Note the two layers — a diffusion layer that permits long-range charge interactions and a CE layer that localizes the charge distribution from the diffusion layer and produces improved SNR via the feedback pathways. Currently, the electrical model is simulated by integrating the governing equations shown in Figure 10. A separate input layer of feature-sensitive neurons provides activity to the diffusion layer via glial cell endfeet. Glial endfeet also bidirectionally carry feedforward and feedback activity (i.e., charge or K^+ ions).

The diffusion layer of the DEB model is governed by three coupled systems of differential equations based on Ohm's law; the first represents the spatially coupled diffusion layer:

$$\begin{aligned}\frac{dQ_g^{(i)}}{dt} &= \frac{G_{gg}}{C_g}\left[Q_g^{(i+1)} + Q_g^{(i-1)} - 2Q_g^{(i)}\right] - \frac{G_g}{C_g}Q_g^{(i)} \\ &\quad + \frac{G_{ge}}{C_g}\left[\frac{C_g}{C_e}Q_e^{(i)} - Q_g^{(i)}\right] + \frac{G_{gi}}{C_g}\left[\frac{C_g}{C_i}Q_i^{(i)} - Q_g^{(i)}\right] \quad . \end{aligned} \tag{1}$$

This equation contains several parameters that can be considered independent of the other coupled equations. Conductivity G_{gg} controls the speed with which charge Q_g is distributed throughout the glial layer, while G_g controls how rapidly charge leaks from the glial nodes into the environment. Together, conductivities G_{gg} and G_g determine the spatial extent over which charge can spread in

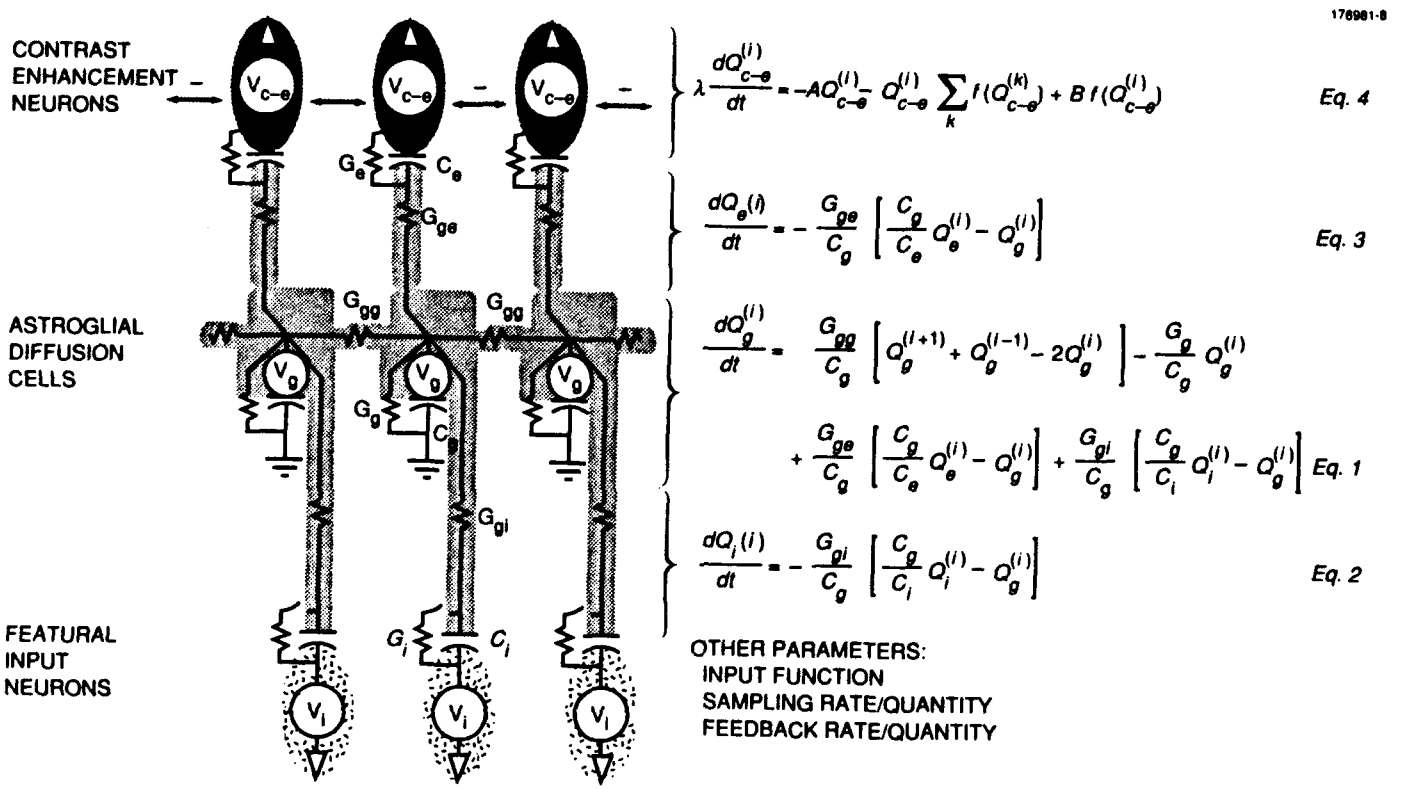


Figure 10. Diffusion-enhancement bilayer (DEB) circuit model. Charge is distributed throughout the network via the dynamics in the listed equations. Note the two conceptual layers (and the input) that provide long-range communication (astroglial diffusion cells) and localization (contrast-enhancement neurons).

the diffusion layer. The other two equations govern charge input to the diffusion layer as provided by the featural input neurons [Equation (2)] and feedback from the CE neurons [Equation (3)].

$$\frac{dQ_i(i)}{dt} = -\frac{G_{gi}}{C_g}\left[\frac{C_g}{C_i}Q_i^{(i)} - Q_g^{(i)}\right] \quad , \tag{2}$$

$$\frac{dQ_e(i)}{dt} = -\frac{G_{ge}}{C_g}\left[\frac{C_g}{C_e}Q_e^{(i)} - Q_g^{(i)}\right] \quad . \tag{3}$$

Conductance G_{gi} controls the rate at which new inputs affect the charge on the diffusion layer. G_{ge} controls the rate at which the CE layer feels the effects of developing charge distributions on the diffusion layer, as well as the rate at which feedback from the CE layer modifies the diffusion layer. The capacitors represented in all three equations store the distribution of charge in the diffusion layer C_g and at the interfaces to the input C_i and the enhancement layers (C_e).

The charge in the CE endfeet is periodically sampled[3] by the CE neurons, which process activity on a shorter time scale than the diffusion layer. The sampled charge is contrast enhanced via a network originally formulated by Grossberg [8], and the output from this network is fed back to the facing endfeet. The equation governing charge in this system of N neurons represents a network of self-exciting nodes with long-range lateral inhibition and passive decay;

$$\lambda\frac{dQ_{c-e}^{(i)}}{dt} = -AQ_{c-e}^{(i)} - Q_{c-e}^{(i)}\sum_k f(Q_{c-e}^{(k)}) + Bf(Q_{c-e}^{(i)}) \quad , \tag{4}$$

where

$$f(Q) = \begin{cases} CQ^2/Q_1 & 0 \le Q \le Q_1 \\ CQ & Q_1 < Q \end{cases} \quad . \tag{5}$$

Equation (4) can be rewritten as a shunting short-term memory model with charge limited to the range [0,B]. Depending on the choice of parameters, the rapidly attained equilibrium can either pick the node with maximal charge or contrast-enhance the charge across the layer. The latter properties are of interest because constant signals are suppressed, noise fluctuations are quenched, and all nodes of nearly maximal activity are enhanced. In any case, the dynamics lead

[3]Periodic sampling in time by the CE neurons can be identified with the refractory period of neurons that are phase-locked in a layer.

to a normalization of activity across the layer, with total equilibrium activity equal to $E = B - \frac{A}{C}$. When in this domain, the nodes for which activities fall below

$$Q(t) = \frac{Q_1}{B - \frac{A}{C}} \tag{6}$$

for a sufficiently large time will be forced to lose all activation, i.e., they will be quenched.

Because feedback reactivates the diffusion layer, even once the original input is off, the feedback amplitude must be dampened over time. Without this step, a single light will be sustained in memory forever. This problem is resolved by forcing the parameter B in Equation (4) to decay with time between inputs to the system; that is, the duration of the feedback is limited in time. When a new input stimulates the visual field, the CE layer is re-energized and B is reset to its maximum value. Between inputs this decay is modeled as

$$B = \Phi\left(B_{\max} - \left(\frac{t}{\tau_1}\right)^2\right) \quad , \tag{7}$$

where $\Phi(x)$ is a threshold linear function equal to x, if $x \geq \phi$, and 0 otherwise. (To date, most experiments have set $\phi = 0$.)

At the end of this report another choice will be motivated for modulating B so that it has a finite rise time, as well as a finite duration. For these experiments, B is chosen as

$$B = B_{\text{scale}} t^p e^{\frac{-t}{\tau_2}} \quad . \tag{8}$$

3.2 Relationship to Biological Networks

A number of DEB model components have direct correlates in biology. Capacitor C_i represents a neuronal-astroglial interconnect that is locally excited by presented features. A feature-sensitive neuron fires; as it repolarizes, K^+ is released into the extracellular compartment. (The input source function described in Section 4.2 tries to model this charge release.) K^+ ion pumps bring K^+ onto C_i, which represents a highly permeable endfoot of an astrocyte glial cell (cf. K^+-spatial buffering [15,18]). Here the K^+ is freely diffused via ion currents within glia (with membrane capacitance C_g) and forms a network through electrical gap junctions between astrocyte glial cells [12]. The interglial connections are represented by the conductors G_{gg}. A portion of the K^+ is diffused out of the cells at endfeet to an upper CE neuronal layer, which is excited by the locally increased extracellular K^+ concentration. This astroglial-neuronal interconnect is represented by C_e. The neuronal layer is hypothesized to interact within itself to contrast-enhance its own activity, further releasing K^+ as it fires. It then feeds this contrast-enhanced K^+ profile back to the glial layer via the same endfeet, thereby reinforcing the charge distribution in the glial network, particularly

near the charge maximum. The output of the CE layer also provides the basis for the percept of a compact form in smooth motion.

3.3 Alternative Models

One interesting alternative model for long-range apparent motion was proposed by Grossberg and Rudd [9]. Their basic model elements responsible for creating continuous motion paths from spatially disparate inputs are very similar to those currently being studied. Essentially, localized inputs (e.g., flashes of light) are assumed to excite a spatially extended Gaussian activation pattern of fixed scale. By combining a preprocessing stage, which detects spatial gradients of brightness with a temporal change detector, their input functions grow and decay over time. When this growth function is used to excite the Gaussian activity pattern, a fixed-scale Gaussian activity wave is obtained, with amplitude that grows then decays in time. Grossberg and Rudd demonstrate that if spatially separate inputs are flashed at different times and an appropriate scale Gaussian used, the two activity waves will merge into a single activity hump, the maximum of which slides continuously from the position of the initial input to that of the final input. They then assume that a separate contrast-enhancing process localizes this moving maximum.

The DEB model shares the two essential elements of the Grossberg-Rudd model, i.e., a spatially extended response to an input that evolves over time, followed by a CE process that localizes the response; however, where Grossberg and Rudd assume a fixed-scale Gaussian response to an input, the DEB uses a diffusion process that responds with increasing scale as a function of time. Both models seek other, earlier processes to determine the dynamic nature of the input function responsible for exciting the activity profiles that will interact with one another.

Another alternative model, introduced by Waxman, et al. [22], and modified by Fay and Waxman [6], is the short-range motion process, the essential concept of which is the temporal growth and decay of a Gaussian activity wave in response to a transient input. For noninteracting inputs (i.e., a Gaussian wave with scale smaller than the spacing of inputs), this process provides a means to directly extract the speed of moving features. When the features are close to one another the waves interact; in so doing, they interpolate the trajectory between inputs.

4. NUMERICAL SIMULATIONS

4.1 Overview

The following numerical simulations continue to build on work performed previously. Tool flexibility and speed continue to be enhanced. This development effort allowed demonstration of many of the percepts discussed in section 2.1.1. After this success the network was reexamined, considering the effect of noise, and solutions were formulated for any problems encountered. This section discusses the successes achieved in chronological order; except where indicated, the following network parameters are employed: $\frac{G_g}{G_{gg}} = .25, \frac{G_{ge}}{G_{gg}} = .4, \frac{G_{gi}}{G_{gg}} = .003$(dynamic) and $\frac{G_{gi}}{G_{gg}} = .15$ (static), $\frac{C_g}{C_e} = 15, \frac{C_g}{C_i} = 5, X_1 = 23, A = 1, B = 2501, C = 1, \phi = 0, F = 2.25$. In the first series of simulations B is modulated as specified in Equation (7), with $B_{\max} = 2501, \tau_1 = 2.25$, while in the later simulations B is modulated as in Equation (8), with $B_{\text{scale}} = 50, \tau_2 = 25, p = 1.5$.

In all simulations the motion of the CE layer's activity edges and the activity maximum are considered to drive the percept of apparent motion, with the latter defining object location.

4.2 System Input

Two different inputs represent the result of early visual processing performed in each of the two visual pathways. The first is a sustained input used in the static pathway; this input rapidly rises to a maximum value that is sustained until the stimuli are removed. The second is a transient input used in the dynamic pathway; this input rises rapidly, then decays quickly. Although attempts were made to produce results consistent with the psychophysical data base for static and dynamic systems using one copy of the model and one type of input, these attempts failed. The neurophysiological literature was then consulted; it was noted that cat X and Y cells (analogous to primate A and B cells) had significantly different stimulus response characteristics [5]. A functional approximation was then generated to the "average response histogram" recorded by Enroth-Cugell and Robson [5]. The two inputs are pictured in Figure 11. The static input is defined as:

$$I(t) = \begin{cases} 20 * t^{1.1} * e^{\frac{-t}{15}} & 0 \leq t < 15 \\ 20 * 15^{1.1} * e^{\frac{-15}{15}} & 15 \leq t < 100 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

while the dynamic input is defined as:

$$I(t) = \begin{cases} 26 * t^{1.1} * e^{\frac{-t}{20}} & 0 \leq t < 70 \\ 0 & \text{otherwise} \end{cases} . \tag{10}$$

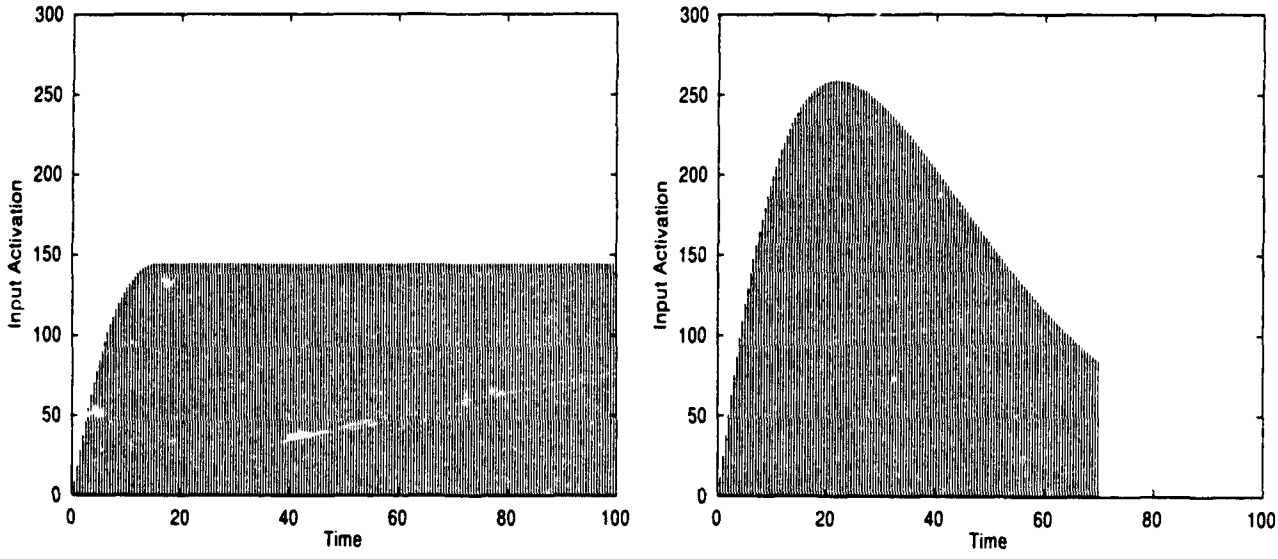


Figure 11. Simulation input. Left is the "sustained" input to the "parvocellular" system, right is the "transient" input to the "magnocellular" system.

4.3 Gamma Motion

In this experiment a single light is illuminated, then extinguished. Recall that the percept is of the light expanding when it is first illuminated, then contracting when it is extinguished. The total activity distributed throughout the system is depicted in Figure 12. (The graphs for this and remaining simulations are presented from left to right, and from top to bottom.) Activity enters the network through the input endfeet (upper left), then flows into the central network where activity is diffused into adjacent glial cells (upper right). This activity distribution further spreads into the glial enhancement endfeet (lower left), from where it is then contrast-enhanced (lower right). In this simulation the activity maximum does not move and thus does not contribute to the percept, but the activity edges expand when the dynamic input is first injected into the DEB network and the activity edges contract when the input is extinguished (see contour plot in Figure 12). It is the motion of the activity edges in the CE layer that yields the motion percept.

4.4 Long-Range Apparent Motion

In this experiment two lights are illuminated at different locations and times; when the timing and object spacing is correct, the percept is of smooth motion from the first light to the second. In this simulation (Figure 13), both the CE layer activity maximum and the activity edges contribute to the motion percept. The maximum is believed to indicate the location of the moving illusory

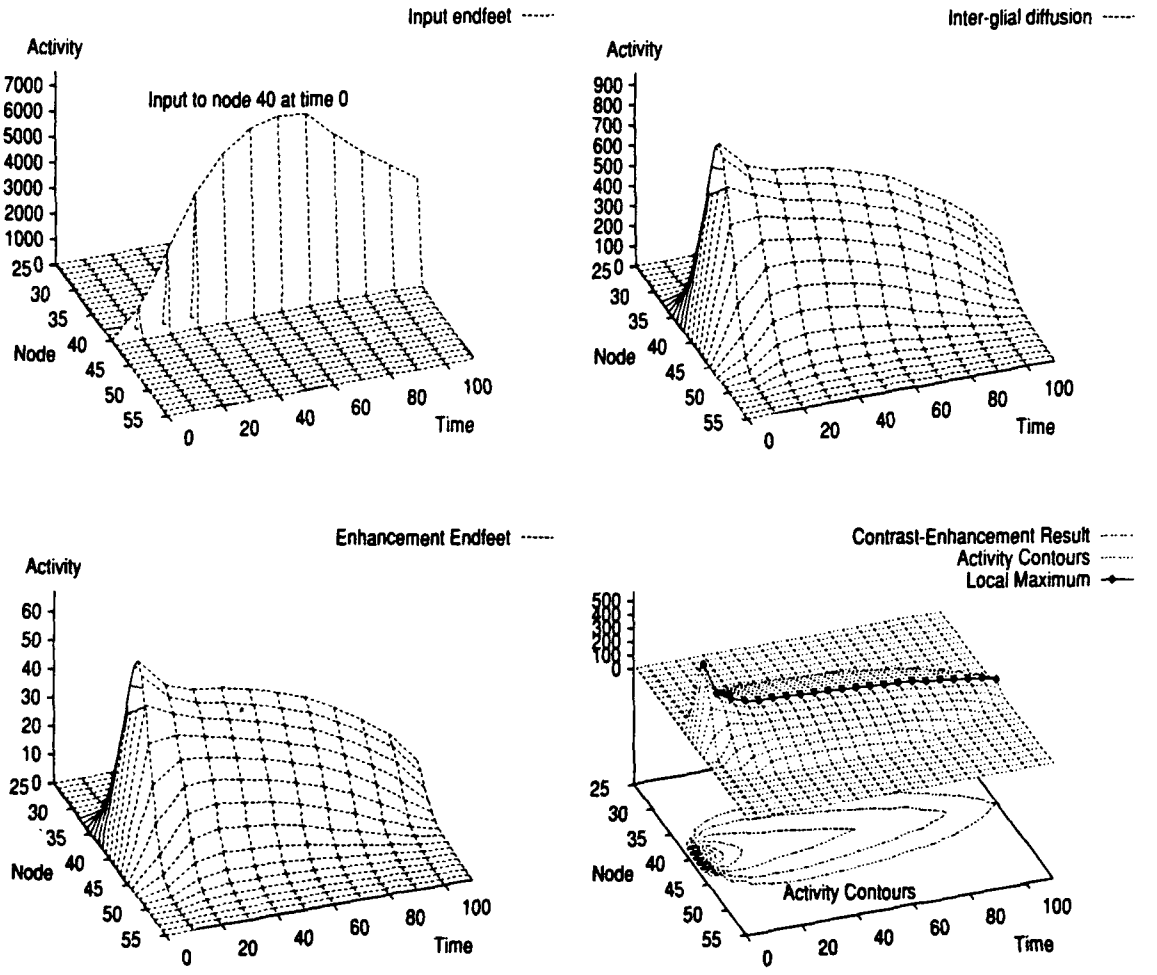


Figure 12. Gamma motion (original feedback). From left to right, top to bottom: activity distributed on glial input endfeet, activity distributed within interglial layer, activity on glial enhancement endfeet, activity after contrast-enhancement processing. The contrast-enhancement activity first expands with stimulus onset, then contracts with stimulus offset. There is no motion of the activity maximum, so this does not contribute to the percept.

light between the two illumination points. Notice how re-energizing the CE layer upon the second input amplifies the activity at the first input. Still the activity maximum moves smoothly between the first and second input locations. The simulation shown here is for stimuli appearing eight nodes apart, but the same simulation can be run (with no parameter changes!) for distances from two to ten nodes away. Slightly greater distances produce incomplete movement from the first to the second stimuli, while still greater distances produce no movement at all. No movement occurs when the activity "mountains" leak from the system before interacting, which is analogous to the non-interacting percept in Figure 3. Incomplete movement occurs when the initial local activity maximum moves toward the distant second maximum, but the distance is great enough that grouping occurs in the manner of the forthcoming "static" grouping examples.

4.5 Equidistant Merge

In this experiment three lights interact; two lights are illuminated, then a third light, equidistant between the first two, is illuminated. The percept is of the first two lights fusing at the third light's location. In the simulation depicted in Figure 14, the three lights' effects can be clearly seen in the input endfeet's activation graph. As the concentrated activity moves from the input endfeet into the interglial communication layer, activity diffuses and begins to interact. This activity moves into the contrast-enhancement endfeet, where the contrast-enhancement layer takes up the activation. Here the merge sensation is reinforced by the motion of activity edges, while the activity maximum moves smoothly from the location of the two initial lights to the central light.[4]

4.6 Equidistant Split

In this experiment three lights interact; one light is illuminated, then two lights, equidistant to the first, are illuminated. The percept is of the first light splitting in two, with each half moving off to join the third light's location. In the simulation depicted in Figure 15, the three lights' effects can be clearly seen in the input endfeet's activation graph. Here the split sensation is reinforced by the motion of activity edges in the CE layer, which expands significantly with the introduction of the two later stimuli. Unfortunately, three final maxima are produced rather than one maximum splitting and joining the later two, so it is difficult to interpret the sensation of object localization. This remains a shortcoming of the model.

4.7 Static Grouping

Figure 16 represents the response of the CE layer to a static 1-D "image" with structure on multiple scales. In this image there are three point-source inputs at nodes 30, 37, and 42.

---

[4]If the third input is removed, a static grouping of two inputs is obtained. Running this simulation in the dynamic pathway produces a broader, weaker maximum than that caused by the merge experiment.

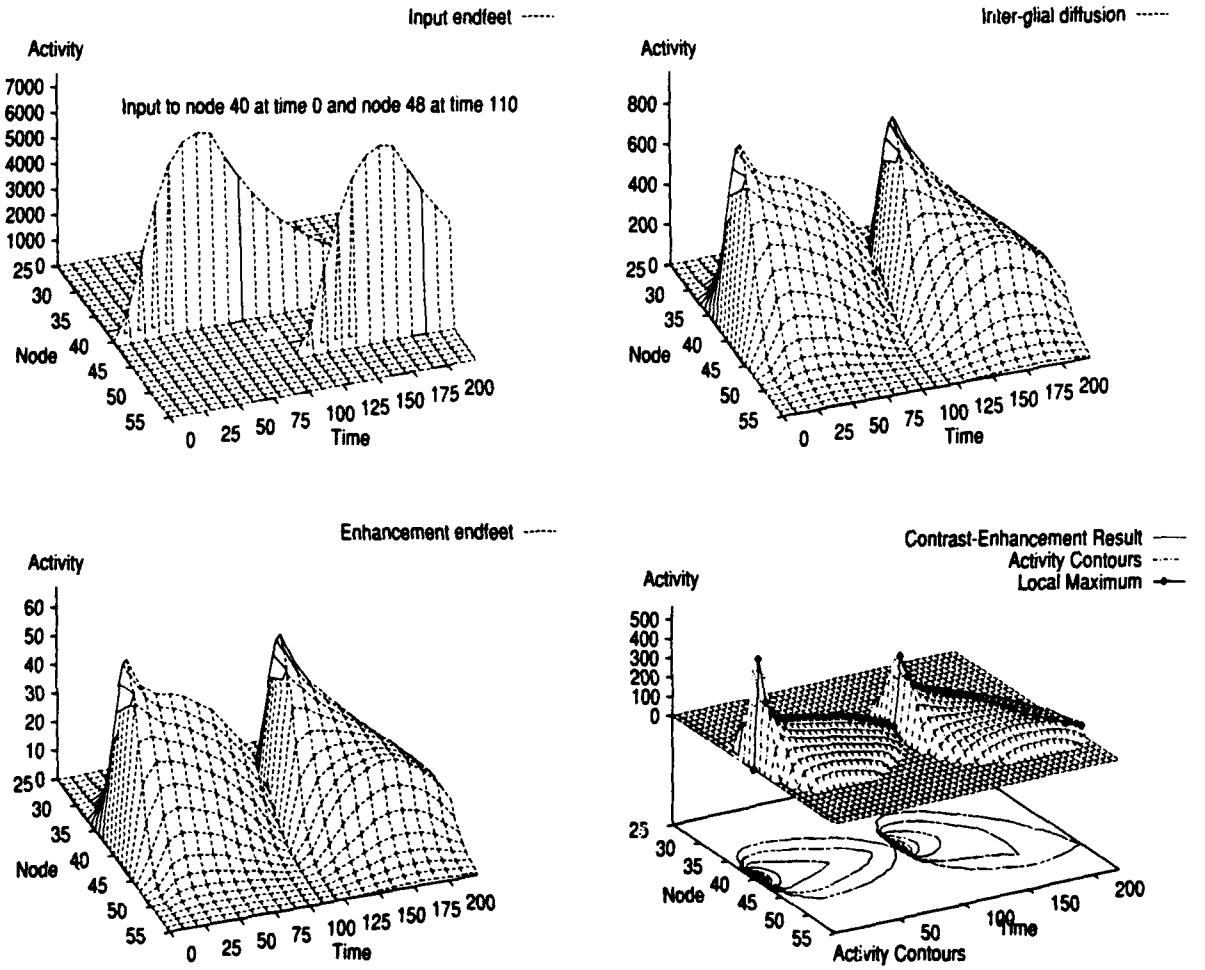


Figure 13. Long-range apparent motion (original feedback). From left to right, top to bottom: activity distributed on glial input endfeet, activity distributed within interglial layer, activity on glial enhancement endfeet, activity after contrast-enhancement processing. Both the contrast-enhanced activity maximum and the activity edges support the percept of motion from node 40 to node 48; the activity maximum provides the location of the object in motion.

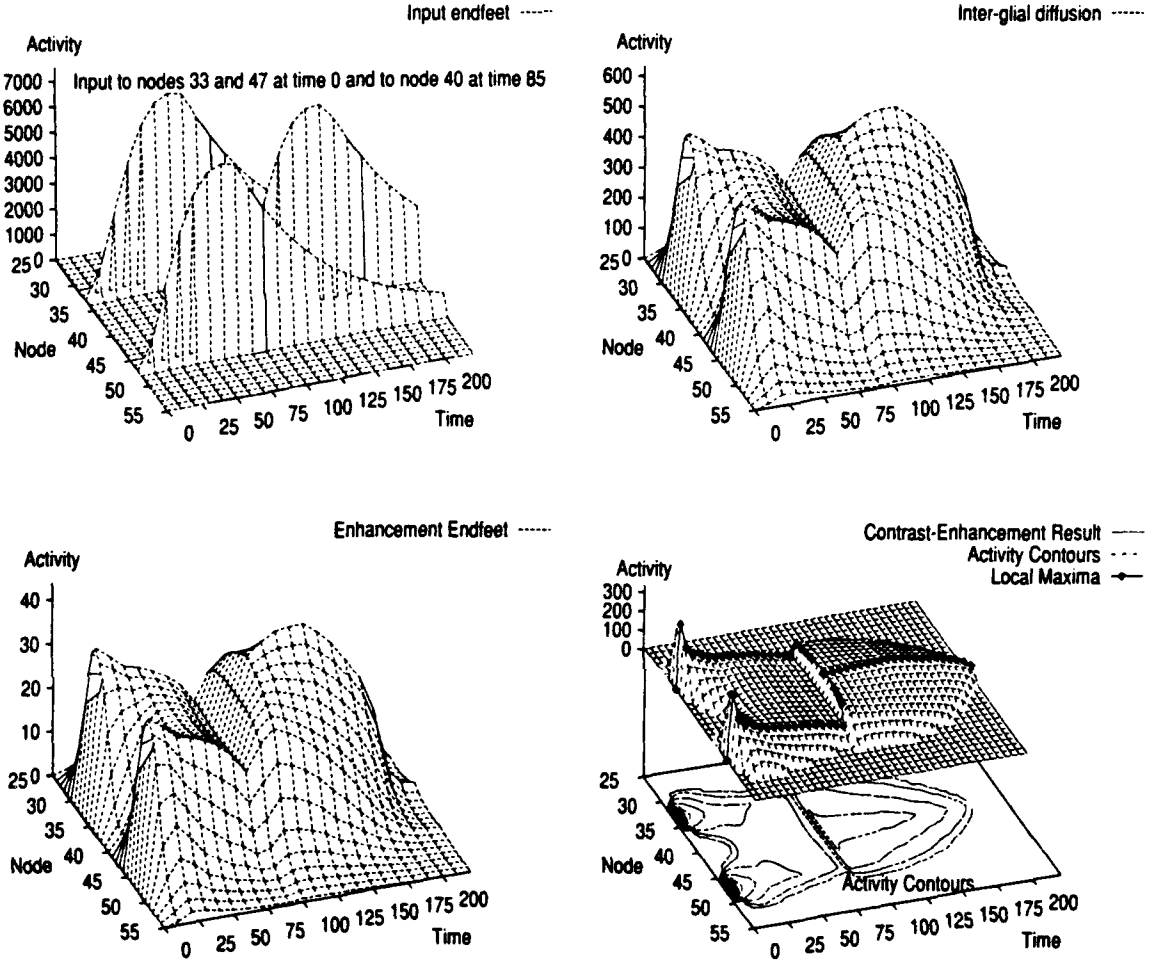


Figure 14. Equidistant merge (original feedback). From left to right, top to bottom: activity distributed on glial input endfeet, activity distributed within interglial layer, activity on glial enhancement endfeet, activity after contrast-enhancement processing. Both the contrast-enhancement's activity maximum and the activity edges support the percept of the two initial stimuli fusing with the central stimulus; the activity maximum provides the location of the stimuli in motion.

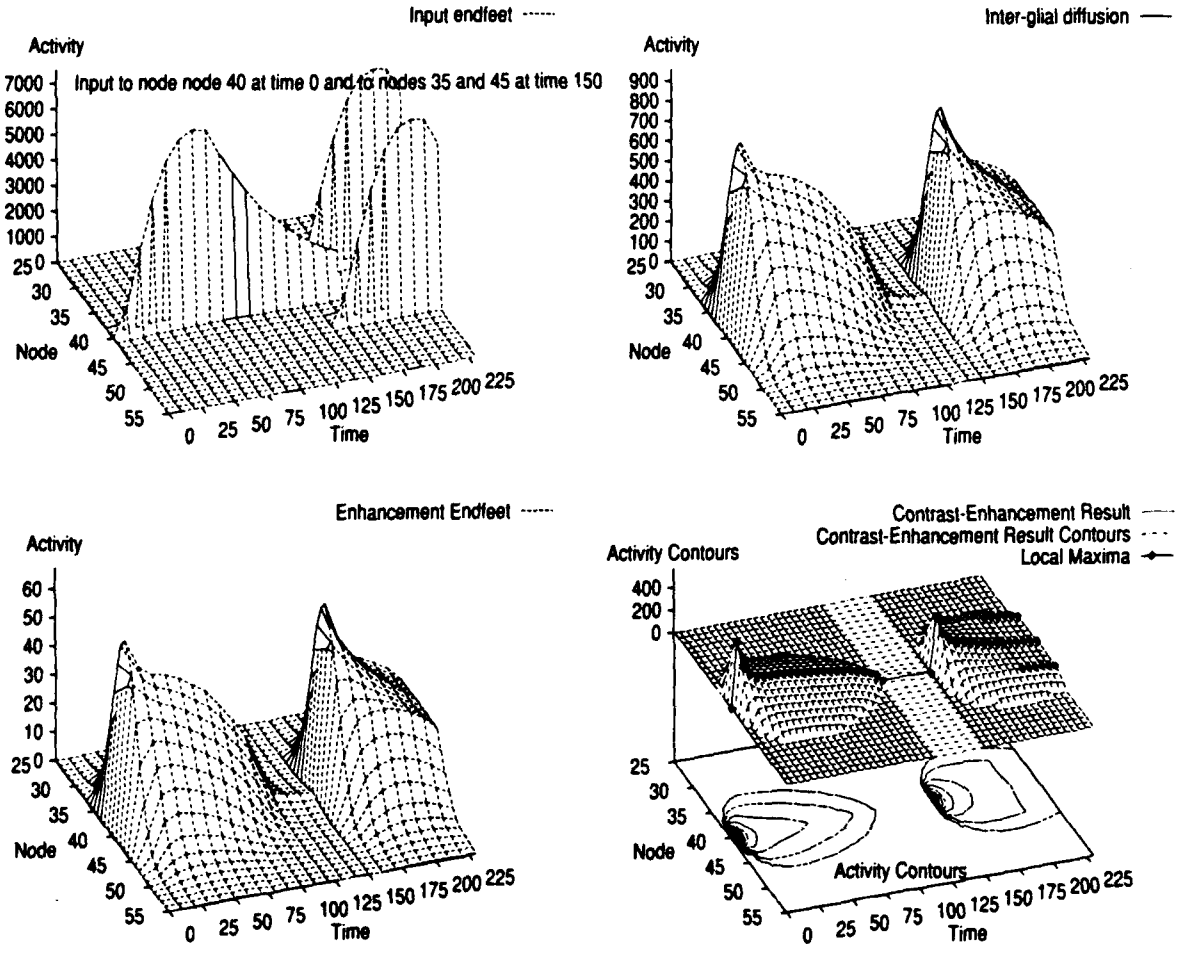


Figure 15. Equidistant split (original feedback). From left to right, top to bottom: activity distributed on glial input endfeet, activity distributed within interglial layer, activity on glial enhancement endfeet, activity after contrast-enhancement processing. Both the contrast-enhanced activity maximum and the activity edges support the percept of the central stimulus fissioning and moving off toward the two introduced stimuli. Three activity maxima result (instead of two, as is perceived), corresponding to the location of the stimuli in motion.

Initially ($t = 1$) the network responds with maxima at each of the three inputs, then ($t = 8$) only two maxima survive as the two closest sources interact and merge on the diffusion layer; finally ($t = 24$), all three merge, and the network displays only a single maximum on this coarser scale. As $t \rightarrow \infty$, this network generates a maximum that can be used as a focus of attention near the geometric mean of the input locations. Note that if the eyes follow the absolute maximum value, concentration is first on the small- then the large-scale interactions.

4.8 Noise Effects

The effect of noise on the system is considered next. In this series of simulations, a small amount of noise is added to the input endfeet, the interglial diffusion sublayer, and the enhancement endfeet of the diffusion layer. Total noise per node is randomly distributed between 0 and 1 unit of activity. This simulates the effects of additive noise that could be caused by residual activity or steady-state random-onset firing of the input neurons. The noise simulations will motivate the choice of a new feedback activity function that has a finite rise time [Equation (8)].

With the original feedback formulation, the gamma motion effect in noise can give rise to false sources of activity (see Figure 17). In this case feedback amplifies small noise fluctuations before the new activity is able to pass through the diffusion layer to the enhancement endfeet. These amplified sources create the appearance of new stimuli where none exist.

By causing the total feedback activity to slowly increase after the stimuli's presence is felt, the stimuli is able to diffuse into the network and overwhelm the noise. One possible formulation that accomplishes this is given in Equation (8), but similar results can be obtained by simply delaying the onset of the feedback activity for a fixed unit of time. With this alternative feedback formulation, the system is less affected by noise and the percept is as expected (see Figure 18). It is interesting to compare the activity contours of Figures 17 and 18; the new formulation expands less rapidly than the old and persists for nearly twice as long (note the change in scale). It is difficult to assess the correctness of either, as no one has yet performed psychophysical experiments to measure the rate of gamma motion expansion.

A similar problem occurs when noise is introduced in the long-range apparent motion simulation and the other simulated effects. In Figure 19, the activity maximum wanders aimlessly due to enhancement of the initial noise.[5]

---

[5]In this diagram, several other weak local maxima exist, but because they are more than 5 orders of magnitude smaller than the depicted local maximum, it is assumed that limited precision neurons would be unable to sense them.

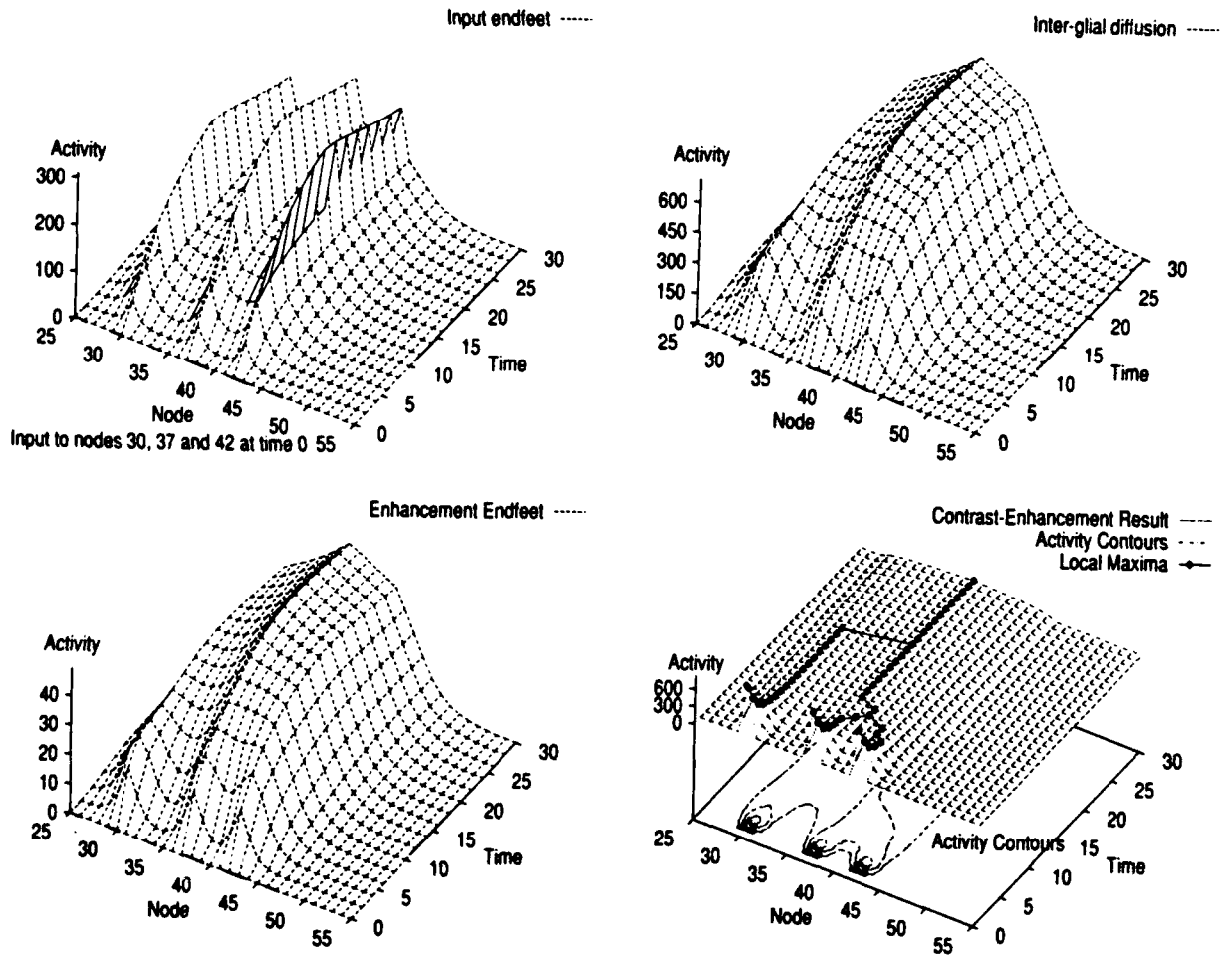


Figure 16. Multiscale static grouping. Three inputs at nodes 30, 37, and 42 interact on increasing scale with increasing time.

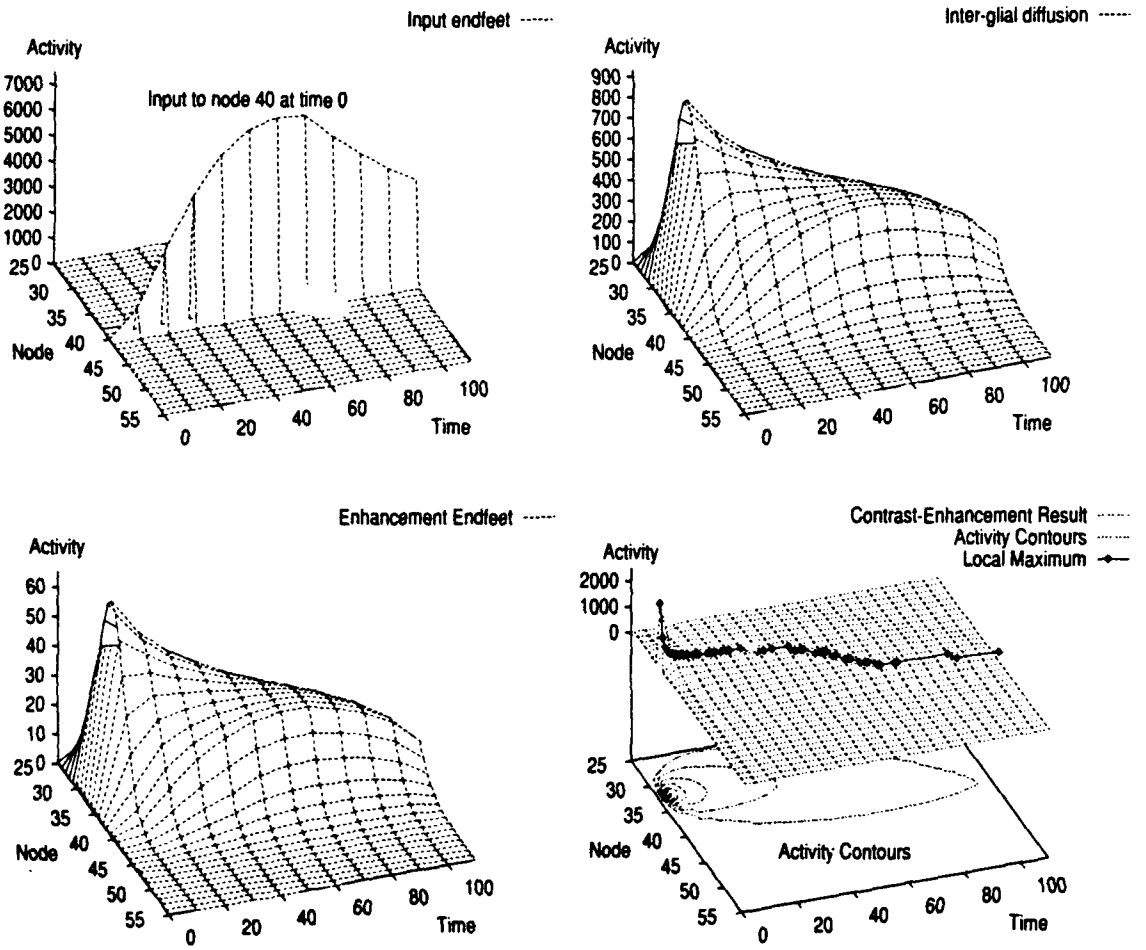


Figure 17. Gamma motion with noise (original feedback). From left to right, top to bottom: activity distributed on glial input endfeet, activity distributed within interglial layer, activity on glial enhancement endfeet, activity after contrast-enhancement processing. Residual activity is enhanced and produces a false long-range apparent motion effect.

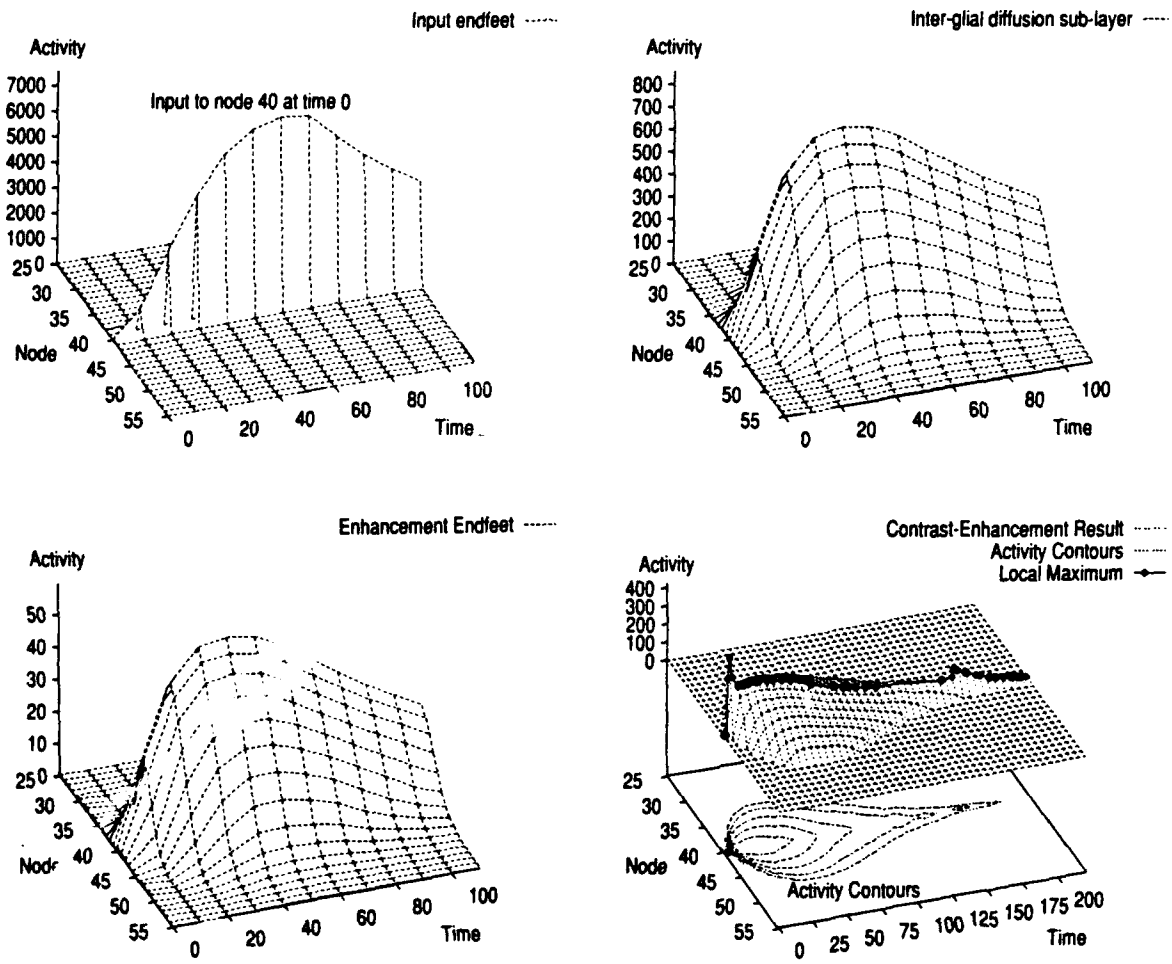


Figure 18. Gamma motion with noise (new feedback). From left to right, top to bottom: activity distributed on glial input endfeet, activity distributed within interglial layer, activity on glial enhancement endfeet, activity after contrast-enhancement processing. The contrast-enhanced activity first expands with stimulus onset and then contracts with stimulus offset. There is no motion of the activity maximum, so this does not contribute to the percept.

In the simulation employing the new feedback (Figure 20), anomalous behavior only appears at the very beginning of the simulation (two maxima occur). One time unit later, the spurious maximum ceases to exist and the simulation then proceeds as it should.[6]

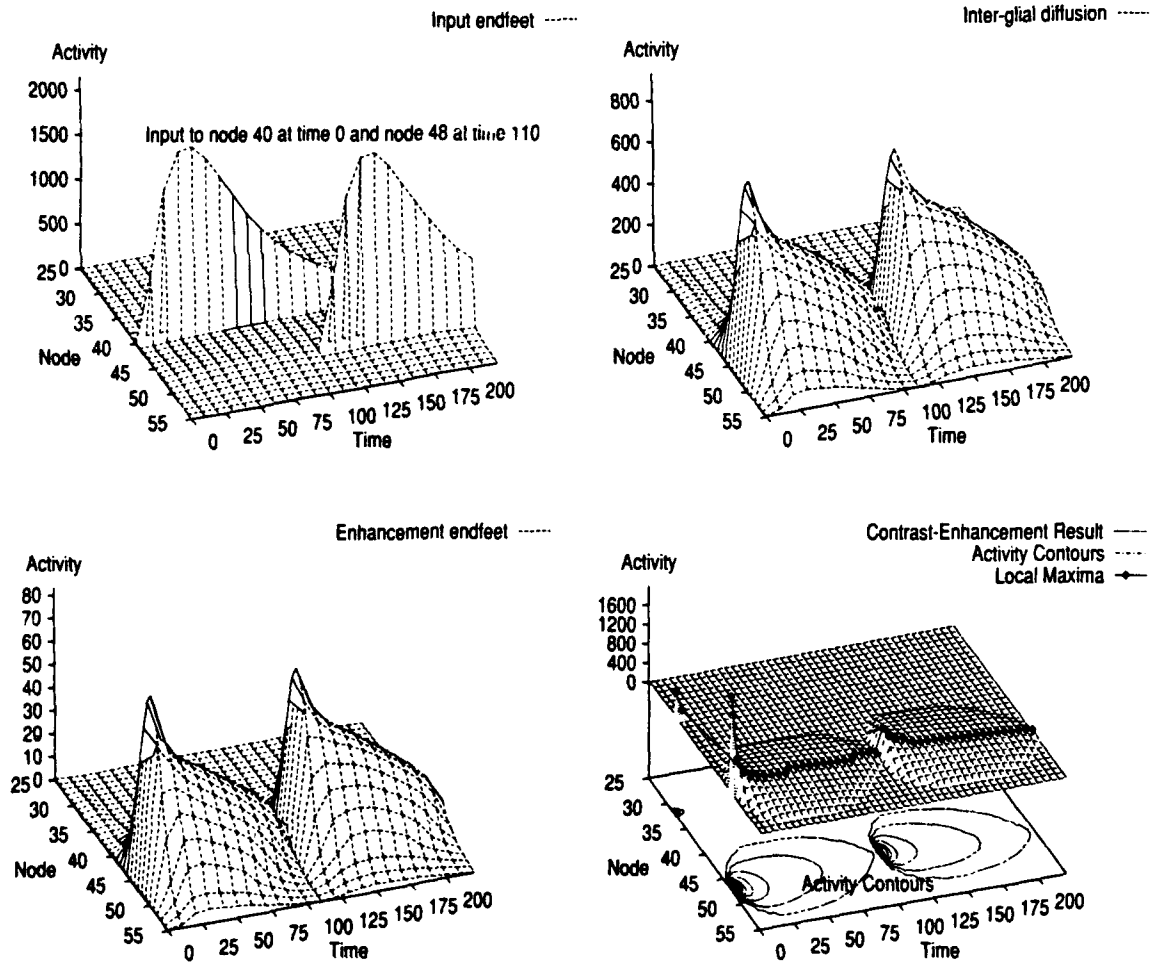


Figure 19. Long-range apparent motion with noise (original feedback). From left to right, top to bottom: activity distributed on glial input endfeet, activity distributed within interglial layer, activity on glial enhancement endfeet, activity after contrast-enhancement processing. Residual activity is enhanced and causes erratic motion.

[6]In this simulation the extremely weak maxima also occur.

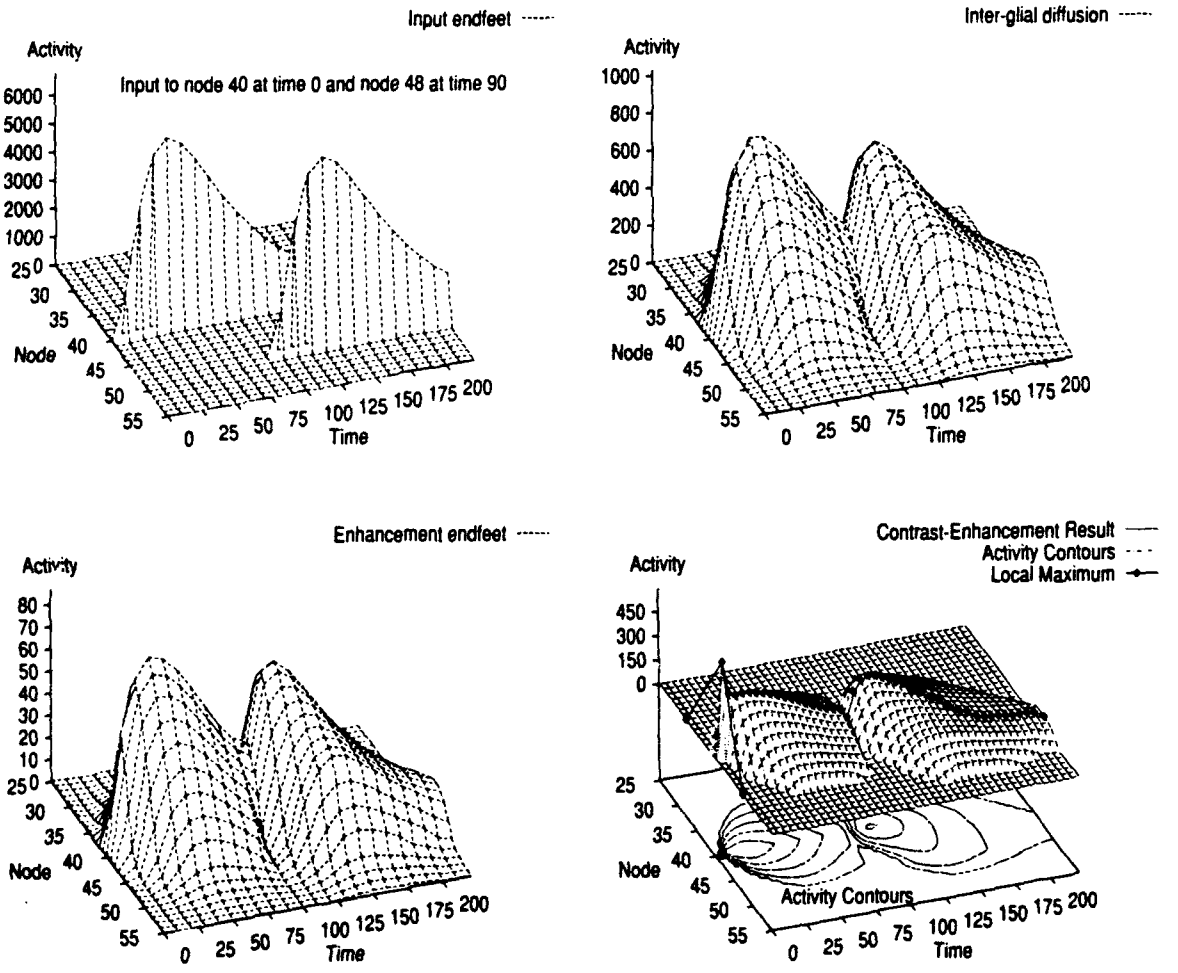


Figure 20. Long-range apparent motion with noise (new feedback). From left to right, top to bottom: activity distributed on glial input endfeet, activity distributed within interglial layer, activity on glial enhancement endfeet, activity after contrast-enhancement processing. Long-range apparent motion occurs, with smooth motion from the first input node to the second input node.

5. WORK IN PROGRESS

A dynamic model for a two-layered network that spatiotemporally groups its inputs on multiple scales as a function of time has been explored. Moreover, this model generates long-range apparent motion between spatially separate inputs introduced at different times. Subtle changes to the rate that feedback is introduced to the system produced a network that was more tolerant to noise.

Current research includes examining the new feedback profile on simulations of equidistant and non-equidistant split, equidistant and non-equidistant merge, and in locating a parameter pool that will produce long-range apparent motion over a wider range of distances. Also, the robustness of the current feedback's formulation will be explored by systematically increasing the noise present in the system. A related series of experiments will explore intersimulation time, to see how long lingering activity will influence a later simulation.

Future work will include the psychophysical phenomena not yet discussed: the Ternus and peripheral effects. The peripheral effects should be simple to implement, as these are believed to be due to the topological preprocessing performed before the data arrives at the network. To simulate the biological system, the data will be spatially logarithmically mapped before administering activity to the system; this procedure is expected to reproduce the peripheral psychophysical results. The output of this system will also be extended via a second layer of *local* contrast-enhancement to more clearly illustrate a pathway concentrating on the expanding activity edges.

The final effort to be explored is the 1-D Ternus effect; when its results can be reproduced, the DEB model will be extended to two dimensions.

REFERENCES

1. J.A. Conners, L.S. Benardo, and D.A. Prince, "Carbon dioxide sensitivity of dye coupling among glia and neurons of the neocortex," *Journal of Neuroscience* 4, 1324–1330 (1984).

2. R.K. Cunningham and A.M. Waxman, "Astroglial-neural networks, diffusion-enhancement bilayers, and spatio-temporal grouping dynamics," in *International Joint Conference on Neural Networks* II, INNS/IEEE, Seattle, WA (1991), p. 937.

3. R.K. Cunningham and A.M. Waxman, "Astroglial-neural networks, diffusion-enhancement bilayers, and spatio-temporal grouping dynamics," in *Sensor Fusion IV: Control paradigms and data structures* 1611 (1991), pp. 411–422.

4. R.K. Cunningham and A.M. Waxman, "Parametric study of diffusion-enhancement networks for spatiotemporal grouping in real-time artificial vision," Annual Technical Summary ESD-TR-91-108, MIT Lincoln Laboratory, Lexington, MA (1991). DTIC-AD-A239674.

5. C. Enroth-Cugell and J.G. Robson, "The contrast sensitivity of retinal ganglion cells of the cat," *Journal of Physiology* 187, 517–552 (1966).

6. D. Fay and A.M. Waxman, "Neurodynamics of real-time image velocity extraction," in *Neural networks for vision and image processing*, Cambridge, MA: MIT Press (1992).

7. F.A. Geldard and C.E. Sherrick, "Space, time and touch," *Scientific American* (April 1986).

8. S. Grossberg, "Contour enhancement, short-term memory, and constancies in reverberating neural networks," *Studies in Applied Mathematics* 52, 217–257 (1973).

9. S. Grossberg and M.E. Rudd, "A neural architecture for visual motion perception: Group and element apparent motion," *Neural Networks* 2, 421–450 (1989).

10. L. Hertz, "Possible role of neuroglia: A potassium-mediated neuronal-neuroglial-neuronal impulse transmission system," *Nature* 4989, 1091–1094 (1965).

11. D.H. Hubel, *Eye, Brain and Vision*, Scientific American Library (1988).

12. H. Kettenman and B.E. Ransom, "Electrical coupling between astrocytes and between oligodendrocytes studied in mammalian cell cultures," *Glia* 1, 64–73 (1988).

13. H.K. Kimelberg and M.D. Norenberg, "Astrocytes," *Scientific American*, pp. 91–96 (April 1989).

14. P.A. Kolers, *Aspects of Motion Perception*, New York: Pergamon Press (1972).

15. S.W. Kuffler, J.G. Nicholls, and R.A. Martin, *From Neuron to Brain*, Sinauer Associates, Inc. (1985).

16. M. Livingstone and D. Hubel, "Segregation of form, color, movement, and depth: Anatomy, physiology, and perception," *Science* 240, 740–749 (1988).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	24 July 1992	Annual Technical Summary April 1991 — June 1992

4. TITLE AND SUBTITLE

Parametric Study of Diffusion-Enhancement Networks for Spatiotemporal Grouping in Real-Time Artificial Vision

5. FUNDING NUMBERS

C — F19628-90-0002
PR — 388
PE — 61102F

6. AUTHOR(S)

Robert K. Cunningham and Allen M. Waxman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Lincoln Laboratory, MIT
P.O. Box 73
Lexington, MA 02173-9108

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Office of Scientific Research
Bolling AFB
Washington, DC 20402

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

ESC-TR-92-121

11. SUPPLEMENTARY NOTES

None

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

This is the second Annual Technical Summary of the MIT Lincoln Laboratory parametric study of diffusion-enhancement networks for spatiotemporal grouping in real-time artificial vision. Spatiotemporal grouping phenomena are examined in the context of static and time-varying imagery. Dynamics that exhibit static feature grouping on multiple scales as a function of time, and long-range apparent motion between time-varying inputs, are developed for a biologically plausible diffusion-enhancement layer coupled by feedforward and feedback connections; input is provided by a separate feature-extracting layer. The model is cast as an analog circuit that is realizable in VLSI, the parameters of which are selected to satisfy a psychophysical data base on apparent motion.

14. SUBJECT TERMS

neural networks
astrocyte glial networks
diffusion enhancement
long-range apparent motion
spatiotemporal grouping dynamics
interference suppression

15. NUMBER OF PAGES
52

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR